



# Description d'un système de détection automatique de la sonie des sons du langage

Bernard Teston

## ► To cite this version:

Bernard Teston. Description d'un système de détection automatique de la sonie des sons du langage.  
Travaux Interdisciplinaires sur la Parole et le Langage, 1979, 6, pp.27-46. hal-00285541

**HAL Id: hal-00285541**

**<https://hal.science/hal-00285541>**

Submitted on 27 Jun 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

TRAVAUX DE L'INSTITUT DE PHONETIQUE D'AIX  
Volume 6, 1979. Pp. 27-46

DESCRIPTION D'UN SYSTEME DE DETECTION AUTOMATIQUE DE LA  
SONIE DES SONS DU LANGAGE

BERNARD TESTON  
(INSTITUT DE PHONETIQUE LA 261)

# DESCRIPTION D'UN SYSTEME DE DETECTION AUTOMATIQUE DE LA SONIE DES SONS DU LANGAGE

## RESUME.

Le système que nous décrivons est un détecteur d'intensité des signaux acoustiques du langage auquel sont appliquées différentes pondérations de manière à obtenir une mesure la plus proche possible de la sensation auditive effectivement perçue. Ces pondérations tiennent compte :

- des courbes isosoniques,
- de la répartition de l'énergie du signal dans le spectre,
- de l'effet de masque.

L'appareil est essentiellement constitué par un analyseur de fréquence en temps réel dont les filtres d'analyse ont une progression simulant les bandes critiques de ZWICKER (1957).

Un détecteur de voisement permet de faire la distinction entre les signaux voisés et non voisés. Si le signal est constitué par des sons voisés, le calcul de la sonie est alors effectué au moyen de la méthode proposée par ROSSI (1971). La valeur de l'intensité pondérée du signal de parole est obtenue directement en phones toutes les 10 millisecondes. L'exploitation des résultats peut être réalisée au moyen d'un oscillographe enregistreur sur lequel est visualisée la courbe d'intensité comme avec un intensimètre classique auquel l'appareil se substitue. Il évite ainsi un fastidieux et long travail de pondération manuelle à partir d'une courbe d'intensité objective. Le système est également connecté à un calculateur pour réaliser des traitements particuliers sur les paramètres prosodiques.

L'analyseur du détecteur de sonie peut effectuer des analyses spectrales par octave et par 1/3 d'octave. Il est également possible de pondérer l'intensité et la constante de temps de chaque bande d'analyse.

Il est prévu de faire évoluer l'appareil pour intégrer l'influence de la durée sur la perception des segments acoustiques. Des recherches dans ce sens sont envisagées à court terme dans notre laboratoire, ainsi que des études sur la perception des consonnes non voisées, dont nous comptons également intégrer les résultats pour améliorer le calcul de la sonie de ces éléments.

## A SYSTEM FOR THE AUTOMATIC DETECTION OF LOUDNESS IN LANGUAGE-SOUNDS

### ABSTRACT

The system described here is an intensity detector for acoustic signals of language incorporating various weightings in order to obtain as close as possible an approximation to the auditory sensation actually perceived.

These weightings take into account :

- the isosonic curves,
- the distribution of the signal's energy in the spectrum,
- masking effects.

The apparatus consists basically of a real-time frequency analyser with a filter progression approximating ZWICKER's critical bands (1957).

A voicing detector makes it possible to distinguish voiced and unvoiced signals. When the signal comprises voiced sounds the loudness is calculated by means of the method proposed in ROSSI (1971). The weighted intensity value is obtained directly in phones every 10 ms. The exploitation of the results can be obtained by means of an oscillographic recorder on which the intensity curve is visualised as with a classic intensity-meter which the apparatus replaces. It thus renders unnecessary the long and fastidious task of manual correction of the objective intensity curve. The system is also connected to a computer for particular treatment of prosodic parameters.

The analyser of the loudness detector can carry out spectral analyses by octave and 1/3 octave. It is also possible to give specific weightings for the intensity and time-constant for each band.

Future improvements to the apparatus will take into account the influence of duration on the perception of acoustic segments. Research in this direction is being undertaken in our laboratory as well as a study of the perception of voiceless consonants. These results will also be taken into account in order to improve the estimation of the loudness of these elements.

#### ABSTRACT

The system described here is a voicing detector for acoustic signals of language. It calculates the intensity of the signal by means of the method proposed in ROSSI (1971). The weighted intensity value is obtained directly in phones every 10 ms. The exploitation of the results can be obtained by means of an oscillographic recorder on which the intensity curve is visualised as with a classic intensity-meter which the apparatus replaces. It thus renders unnecessary the long and fastidious task of manual correction of the objective intensity curve. The system is also connected to a computer for particular treatment of prosodic parameters.

# DESCRIPTION D'UN SYSTEME DE DETECTION AUTOMATIQUE DE LA SONIE DES SONS DU LANGAGE

B. TESTON

INSTITUT DE PHONETIQUE D'AIX-EN-PROVENCE

LABORATOIRE ASSOCIE AU C.N.R.S., N° 261

## I. INTRODUCTION

Le système que nous décrivons est un détecteur d'intensité des signaux acoustiques du langage. Il diffère des différents intensimètres bien connus des Phonéticiens par le fait qu'il permet d'obtenir une courbe image des variations de l'intensité des sons du langage corrigée en fonction des lois de la perception psychoacoustique.

Les détecteurs d'intensité généralement employés pour des recherches phonétiques réalisent une mesure objective de l'intensité avec parfois une pondération subjective normalisée que l'on trouve dans les différents sonomètres utilisés pour les mesures acoustiques. Ces pondérations, dérivées des courbes isosoniques sont définies pour des niveaux de 40, 70 et 100 phones et sont baptisées respectivement A, B et C. Les Phonéticiens, lorsqu'ils les utilisent, emploient généralement la courbe B. Les résultats obtenus au moyen de ces pondérations pour la mesure de l'intensité effectivement perçue, ne correspondent pas à la réalité. Ces pondérations ne sont valables que pour des sons purs ou des bruits à bande étroite. Les mécanismes de la perception obéissent également à d'autres lois. La sensation auditive ou sonie, outre les différences de sensibilité en fonction de la fréquence définies par les courbes isosoniques dépend : de la répartition de l'énergie du signal dans le spectre, de l'effet de masque, de la durée des segments acoustiques.

Depuis les travaux de FLETCHER et MUNSON (1933) qui définissent les courbes isosoniques, de nombreux chercheurs ont proposé différentes

méthodes de calcul de la sensation sonore effectivement perçue ou sonie, d'un phénomène acoustique. Parmi eux, NIESE (1965), STEVENS (1956) et ZWICKER (1960) les méthodes de calcul de ces deux derniers ont été normalisées par l'ISO. C'est à partir des travaux de ces auteurs que ROSSI (1971) en étudiant l'intensité spécifique des voyelles a été amené à définir une méthode de calcul de la sonie des voyelles. C'est à la lumière de ces résultats que nous avons envisagé la réalisation d'un détecteur automatique de sonie pour les signaux du langage.

Après avoir fait subir à des sujets deux méthodes de tests différentes, ROSSI (1971) déduit de ses observations psychoacoustiques un facteur de pondération Isp caractéristique des différentes voyelles. L'intensité spécifique peut être évaluée à partir de ce facteur de pondération qui permet de corriger l'intensité objective mesurée afin d'obtenir des voyelles isophones. Puis en essayant de calculer ce facteur de pondération, il utilise successivement les différentes méthodes des auteurs cités en référence et les compare aux valeurs de l'Isp observé. Il apparaît dans ces investigations que la méthode de ZWICKER (1960-61) ne donne pas satisfaction surtout pour les voyelles diffuses à spectre large. La raison semble se situer dans le fait que cette méthode ne rend pas suffisamment compte de la largeur de la bande de fréquence utile. Cependant elle est reconnue comme la plus fondée pour la mesure de la sonie des bruits à faible structure formantique où les pôles d'énergie sont relativement peu importants dans le spectre et ne sont pas séparés par des zéros d'énergie.

La méthode de NIESE (1965) surestime systématiquement l'Isp surtout pour les voyelles à spectre large. La méthode de STEVENS (1956) ne donne pas de résultats sensiblement meilleurs. Pour améliorer le calcul de l'Isp, ROSSI (1971) apporte à la méthode de STEVENS (1956), des corrections particulières qui consistent :

a) à réaliser une analyse acoustique selon les monogrammes de ZWICKER (1965)

b) à augmenter l'importance de la largeur de la bande utile des voyelles et de la répartition de leurs formants surtout s'ils sont séparés par des zéros d'énergie (voyelles diffuses) en utilisant la notion de facteur F de STEVENS (1956). Cette procédure peut être résumée ainsi :

1) Après avoir transformé l'analyse en 1/3 d'octave au moyen des monogrammes de ZWICKER (1965), converti les niveaux d'intensité en PHONES au moyen des courbes isosoniques pour sons purs et frontaux, puis en SONES au moyen des tables de STEVENS (1956), on applique la formule suivante :

$$S_{\text{Totale}} \neq S_j + S_k F (/I_k/) \quad k \neq j$$

Cette formule peut être explicitée de la manière suivante :

La sonie totale est égale à la sonie de la Jième bande critique (N° de la bande correspondant à la sonie maximale) augmentée de la somme pour K différent de j, des produits de la sonie de la Kième bande par la fonction F de l'intervalle I, mesuré en octaves, de la Kième bande par rapport au maximum.



- a) la fonction F est définie par le graphe de STEVENS (1956 p.820)
- b) l'intervalle  $I = 3,32 \log \frac{f_K}{f_j}$

La comparaison de l'Is<sub>p</sub> observé et mesuré par les méthodes de NIESE (1965) et STEVENS (1956) corrigé par ROSSI (1971) est donnée dans le tableau I. La méthode de calcul proposée par ROSSI (1971) donne les meilleurs résultats, cependant elle est passablement compliquée à mettre en oeuvre normalement et s'avère difficilement réalisable au moyen d'une procédure automatique. Pour les méthodes de calcul normalisées de la sonie, telles les méthodes MARK 7 de STEVENS (1961) et de SWICKER (1960), il existe des programmes de calcul sur ordinateur tels que ceux proposés par PAULUS et SWICKER (1972) ou par les constructeurs d'analyseur de fréquence. Les méthodes originales sont alors plus ou moins simplifiées. Un constructeur a même commercialisé un analyseur automatique de sonie selon ZWICKER pendant un certain temps, mais cet appareil ne semble pas avoir eu un grand succès commercial. (Analyseur HEWLET-PACKARD Type 8.051 A). Nous avons, pour le calcul de la sonie des voyelles légèrement simplifié la méthode de ROSSI (1971) et pour le calcul de la sonie des bruits, nous avons appliqué presque intégralement la méthode de ZWICKER (1960).

### 3. DESCRIPTION DU DETECTEUR.

Le détecteur de sonie<sup>est</sup> essentiellement constitué par un analyseur de spectre en temps réel, un détecteur d'intensité objectif, un détecteur de voisement et différents circuits qui permettent de corriger la courbe d'intensité objective. Son schéma de principe est donné à la figure I.

#### A) L'analyseur de spectre :

Il est constitué par 25 filtres passe-bande en parallèle simulants les bandes critiques définies par ZWICKER (1957, p. 557) tableau N° 2. Ces filtres sont réalisés par la combinaison de filtres actifs passe-haut et passe-bas de BUTTERWORTH selon la méthode de synthèse de SALLEN et KEY (1955). Ils sont caractérisés par une très bonne définition du gabarit, une parfaite linéarité dans la bande transmise et un minimum de temps de propagation compte tenu de leur largeur de bande. Leur pente d'atténuation répond à la recommandation de l'A.N.S.I. CLASS 3 pour les filtres 1/3 d'octave (60 dB par octave). Chaque filtre est associé à un détecteur de valeur efficace d'un facteur de crête de 4 dont la constante de temps peut varier.

Avant la détection, nous appliquons une pondération de l'intensité dans la bande critique en fonction des courbes isosoniques pour sons purs et frontaux (I.S.O. N.F. 30.003) tableau 2. Pour les bandes critiques où l'intensité suit la loi logarithmique, nous utilisons une simple atténuation calibrée, dans le cas contraire, la fonction de pondération comprise entre une racine carrée et un logarithme est réalisée au moyen d'un circuit particulier.

Les signaux, image de l'intensité dans les différentes bandes, corrigée selon les courbes isosoniques, sont ensuite envoyés dans un multiplexeur.

La réalisation d'un tel analyseur représente un investissement important en argent et en travail. Pour cela, nous avons adjoint à l'analyseur des possibilités supplémentaires qui, si elles ne servent pas directement pour le calcul de la sonie, permettent de réaliser des manipulations expérimentales différentes. L'analyseur peut ainsi se transformer en analyseur par octave et par 1/3 d'octave. L'intensité dans chaque bande peut être pondérée au moyen d'un atténuateur de 59 dB par bonds de 1 dB, un sommateur général permet de restituer le signal reconstitué.

B) Le détecteur d'intensité objectif :

Il est essentiellement constitué par un détecteur de valeur efficace capable de mesurer des signaux d'un facteur de crête de 8 avec une distorsion inférieure à 0,5 %. L'intégration est une moyenne pondérée exponentielle dont la constante de temps est variable. Nous utilisons généralement la valeur de 10 millisecondes. La dynamique du détecteur est de 60 dB.

C) Le détecteur de voisement :

Il est relativement rudimentaire et consiste en un détecteur de niveau dans une bande de fréquence particulière. Cette dernière est définie par un filtre passe-bande de 6 dB par octave de chute de part et d'autre de la fréquence de 200 Hertz. Un détecteur de seuil équipé d'une constante de temps décide si le signal est voisé ou non.

D) Les circuits de pondération :

Après multiplexage des signaux issus des filtres et de l'intensimètre objectif, nous effectuons une pondération qui simule l'effet de masque produit sur une bande de fréquence par la bande précédente si son intensité est supérieure.

Cette fonction est réalisée au moyen d'un circuit de décharge exponentielle. Si nous sommes en présence d'un signal non voisé, nous réalisons une sommation de l'énergie des bandes du spectre au-dessus de 40 dB. Cette opération est effectuée toutes les 10 millisecondes.

Si nous devons pondérer un signal voisé, nous effectuons la différence de l'intensité globale avec celle de la bande contenant le maximum d'énergie. Cette quantité, nous fournit une première approximation de la largeur de bande. Un circuit particulier permet d'obtenir le nombre de bandes critiques utiles ainsi que de détecter, s'il



en existe, des zéros d'énergie entre les formants. Ces différentes procédures permettent d'obtenir toutes les 10 millisecondes un coefficient de pondération qui agit sur la modulation du signal d'intensité objectif.

#### 4) RESULTATS

Nous obtenons actuellement à partir d'un spectre 1/3 d'octave connu les résultats exposés au tableau I dans le calcul de l'Isp. Ils sont relativement cohérente avec les résultats de l'observation. Ils sont meilleurs que l'Isp calculé par la méthode de NIESE (1965) et légèrement plus fluctuants que ceux calculés au moyen de la méthode de ROSSI (1971). Nous espérons dans l'immédiat pouvoir affiner le calcul automatique de l'Isp des voyelles. La sonie des consonnes non voisée nous pose par contre un problème car nous ne connaissons pas de résultats particuliers à ce type de segments acoustiques. Nous avons donc fixé arbitrairement le niveau de sonie maximal à 5 db au-dessous de la sonie de la voyelle "a". Une étude sur la perception des consonnes vient de commencer dans notre laboratoire et nous ne tarderons pas à en appliquer les résultats.

#### 5) EXPLOITATION DES RESULTATS

La courbe d'intensité pondérée automatiquement peut être visualisée au moyen d'un enregistreur galvanométrique comme avec un intensimètre classique. Elle est également disponible sous forme de signaux numériques ; il en est de même des niveaux d'énergie dans les bandes d'analyse. Ceci permet un traitement ultérieur sur ordinateur pour l'étude des paramètres prosodiques (TESTON et ROSSI 1977) ou des spectres acoustiques. Ces derniers sont visualisés sur un scope et peuvent être mis en mémoire et restitués sur une table traçante ; il est également possible de réaliser des spectres moyens et des spectres de valeur maximale comme sur la plupart des analyseurs de spectre en temps réel.

#### 6) CONCLUSION

Notre système n'a pas la prétention de faire des miracles, c'est avant tout un appareil conçu pour décharger au maximum les chercheurs des manipulations longues et compliquées. Il est fait pour donner de grandes quantités de mesures sur de nombreux et volumineux corpus de manière commode. Cependant, il est conçu pour pouvoir évoluer très vite en fonction des connaissances futures sur la perception du langage. Il est d'ailleurs en cours d'évolution actuellement pour une meilleure définition des filtres d'analyse (disymétrie des bandes critiques).

Il pourra bientôt tenir compte de l'influence de la durée des segments vocaliques que nous savons très importants dans le calcul de la sonie, mais que nous n'avons pas encore appliquée dans une détection automatique (ZWICKER 1977).

Un programme de recherche sur ce sujet est également en cours dans notre laboratoire et nous espérons pouvoir utiliser les résultats de ces travaux à moyen terme. L'intensimètre ne sera plus autonome mais couplé en permanence à un ordinateur pour le calcul de la pondération en fonction de la durée.

## 2) EXPLOITATION DES RESULTATS

La courbe d'intensité pondérée automatiquement peut être visualisée au moyen d'un enregistreur galvanométrique comme avec un intensimètre classique. Elle est également disponible sous forme de signaux numériques ; il est de même des niveaux d'énergie dans les bandes d'analyse. Ceci permet un traitement ultérieur sur calculateurs pour l'étude des paramètres psychoacoustiques (TESTON et ROSSI 1977) ou des spectres acoustiques. Les données sont visualisées sur un scope et peuvent être mis en mémoire et restituées sur une table tracée ; il est également possible de restituer des spectres moyens et des spectres de valeur maximale comme sur la plupart des analyseurs de spectre en temps réel.

## 3) CONCLUSION

Notre système a été la première tentative de faire des mesures, c'est

voyelle	Isp observé dB	Isp calculé NIESE 1965	Isp calculé STEVENS corrigé ROSSI 1971	Isp calculé par le détecteur de Sonie
a	0	0	0	0
ɔ̃	0,75	2,5	0,6	0,5
ɔ	1	2,5	0,6	1,5
œ	1,25	2,5	0,8	1,5
ɑ̃	1,25	2,5	1	2
o	1,70	3,5	1,7	2
ɛ	1,75	3,5	1,7	2,5
ɛ	1,75	3,5	2,2	2,5
u	2,35	4	2,5	2,5
y	2,50	6	3	3
e	4	6,5	4,2	3,5
i	4,50	7	5,1	5,5

Tableau N° 1

*Valeurs du facteur de pondération Isp des différentes voyelles qui permet de corriger leur intensité objective afin de les obtenir isophones.*

N° du Filtre	Fréquence centrale Hertz	dB 80 à 70	dB 70 à 60	dB 60 à 50	db 50 à 40	db 40 à 30	dB 30 à 20
1	65	9	8	7	7	8	7
2	155	10	8	9	8	8	8
3	250	10	9	9	9	8	9
4	345	10	10	9	9	8	9
5	450	10	10	9	9	9	9
6	560	11	10	9	10	9	9
7	690	10	10	10	10	9	10
8	830	10	10	10	10	10	9
9	980	10	10	10	10	10	10
10	1.155	10	10	10	10	10	10
11	1.355	10	10	10	10	10	10
12	1.580	10	10	10	10	10	10
13	1.835	10	10	10	10	10	10
14	2.130	9	10	10	9	10	11
15	2.480	9	10	10	10	10	10
16	2.900	9	9	9	10	10	10
17	3.400	9	9	10	9	10	10
18	4.020	9	9	10	9	10	10
19	4.780	10	9	10	10	10	10
20	5.700	10	9	9	9	9	9
21	6.850	10	10	10	9	9	9
22	8.400	10	9	9	9	9	9
23	10.500	10	10	10	8	9	8
24	13.300	11	9	9	9	8	8
25	17.300	11	8	8	8	7	5

Tableau N° 2

Valeurs en PHONES des différents niveaux d'intensité dans les bandes critiques dans la gamme comprise entre 20 et 80 dB. Les valeurs sont déduites des courbes isosoniques (N.F. 30.003).

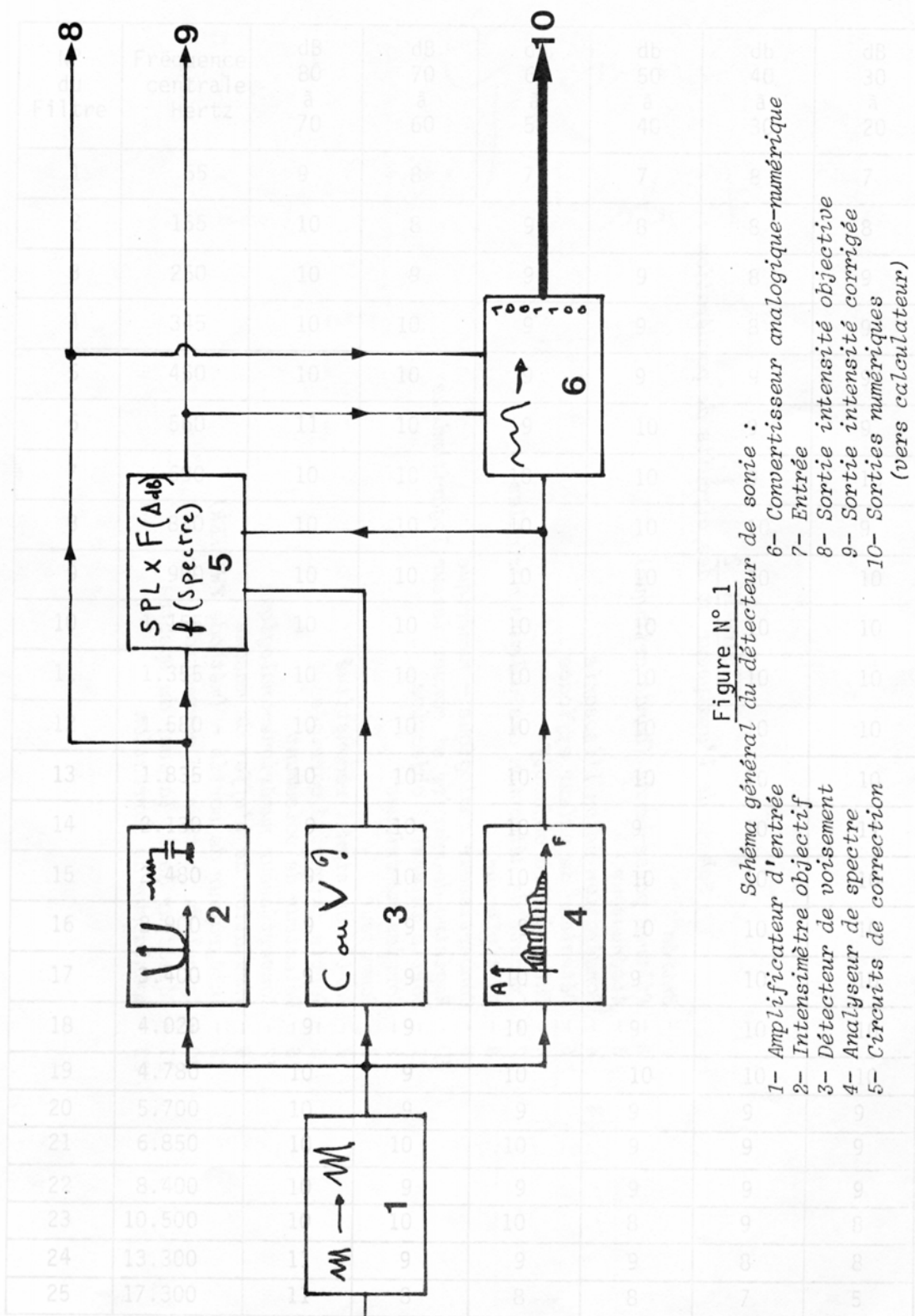


Figure N° 1 Schéma général du détecteur de sonie :

- 1- Amplificateur d'entrée
- 2- Intensimètre objectif
- 3- Détecteur de voisement
- 4- Analyseur de spectre
- 5- Circuits de correction

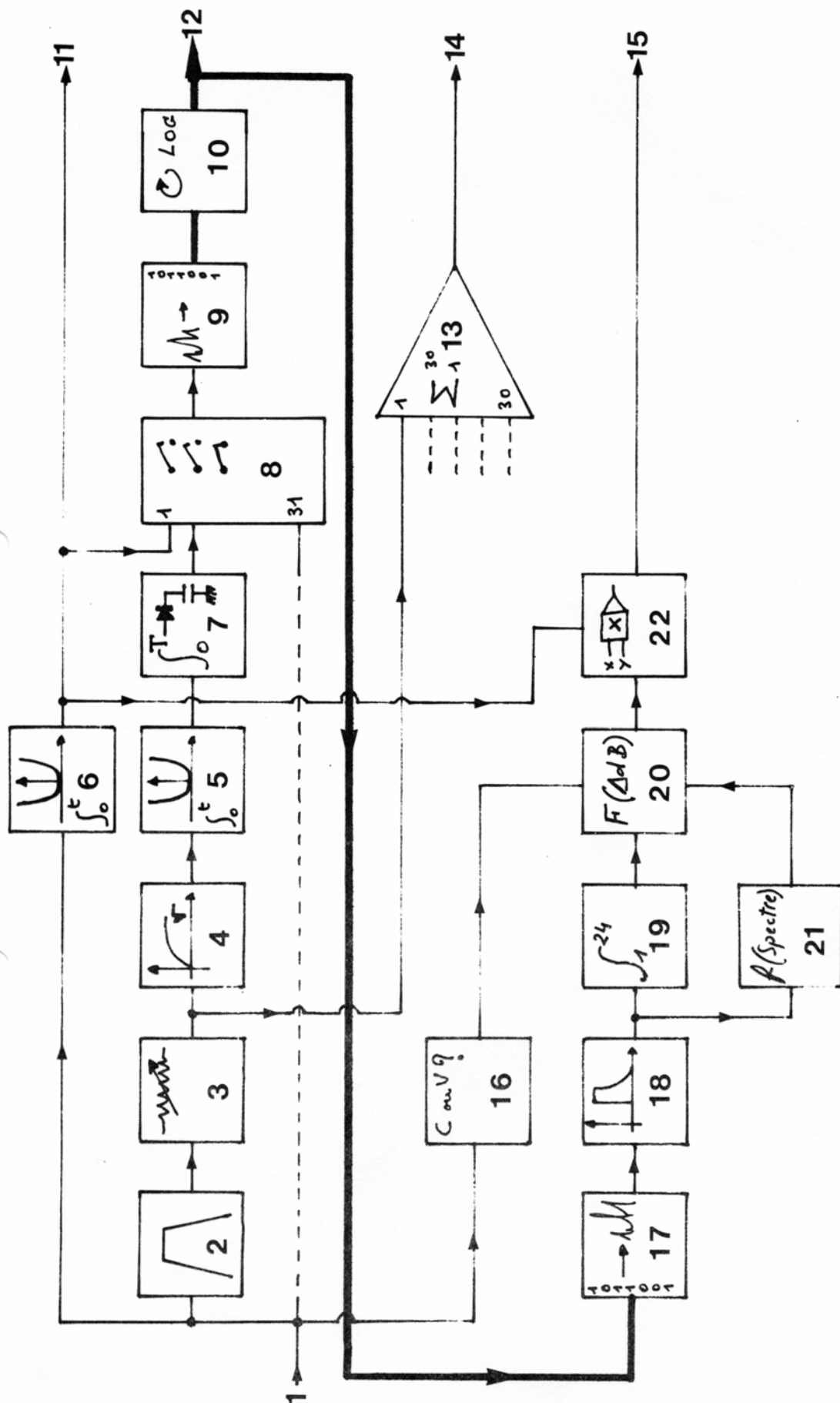
- 6- Convertisseur analogique-numérique
- 7- Entrée
- 8- Sortie intensité objective
- 9- Sortie intensité corrigée
- 10- Sorties numériques (vers calculateur)

Tableau N° 2

Valeurs en PHONES des différents niveaux d'intensité dans les bandes critiques dans la gamme comprise entre 20 et 80 dB. Les valeurs sont déduites des courbes isosoniques (N.F. 30.003).







BIBLIOGRAPHIE

- FLETCHER, H. et MUNSON, W.A. (1933)  
"Loudness, its definition, measurement and calculation",  
JASA 5, 82-108.
- NIESE, H. (1965)  
"Eine methode zur bestimmung der lautstärke beliebiger geräusche",  
ACUSTICA 15, 117-126.
- PAULUS, E. et ZWICKER, E. (1972)  
"Programme zur automatischen bestimmung der lautheit aus terzpegeln  
oder frequenzgruppenpegeln", ACUSTICA 25, 253-266.
- ROSSI, M. (1971)  
"L'intensité spécifique des voyelles", PHONETICA 24, 129-161.
- SALLÉN, R.P. et KEY, E.L. (1955)  
"A practical method of designing RC active filters",  
IRE CT MARCH, 74-85.
- STEVENS, S.S. (1971)  
"Perceived level of noise by MARK 7 and decibel E" JASA 51, 575-601.
- TESTON, B. et ROSSI, M. (1977)  
"Un système de détection automatique du fondamental et de l'intensité",  
8ème J.E.P. GALF, 111-117.
- ZWICKER, E., FLOTTORP, G. et STEVENS, S.S. (1957)  
"Critical bandwith in loudness summation", JASA 29, 548-557.
- ZWICKER, E. (1960)  
"Ein verfahren zur bertimmung der lautstärke", ACUSTICA 10, 304.
- ZWICKER, E. (1977)  
"Procedure for calculating Loudness for temporally variable sounds",  
JASA 62, 675-682.